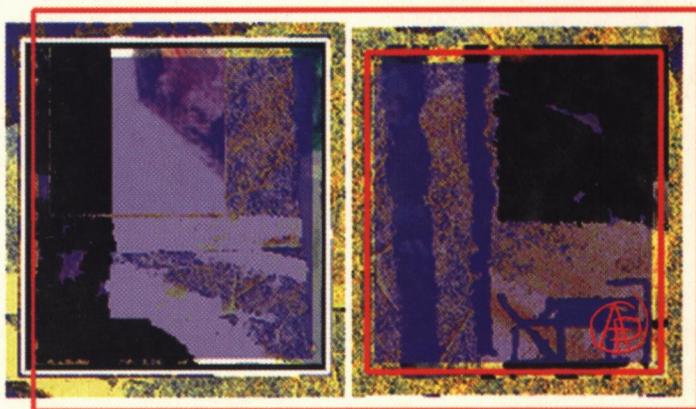


**Sylvie Thiria   Yves Lechevallier**  
**Olivier Gascuel   Stéphane Canu**

*2<sup>e</sup> CYCLE • ÉCOLES D'INGÉNIEURS*

# **Statistique et méthodes neuronales**



**DUNOD**

# TABLE DES MATIÈRES

## Partie 1 : Discrimination

### CHAPITRE 1. RESEAUX DE NEURONES ET STATISTIQUES. UNE INTRODUCTION

FRANÇOISE FOGELMAN SOULIE

<b>1.1</b>	<b>Introduction</b>	1
<b>1.2</b>	<b>Le problème</b>	3
1.2.1	Le modèle statistique	3
1.2.2	L'estimation	3
1.2.3	Espace de recherche de l'estimateur	5
1.2.4	Le dilemme biais / variance	6
1.2.5	Quelques propriétés des perceptrons multicouches	7
<b>1.3</b>	<b>Le modèle de Vapnik</b>	12
<b>1.4</b>	<b>Pénalisation - régularisation</b>	14
1.4.1	Problème mal posé	14
1.4.2	Quelques régularisateurs utilisés dans les réseaux de neurones	15
1.4.3	Le formalisme bayésien	15
<b>1.5</b>	<b>Des familles de méthodes</b>	17
1.5.1	Les méthodes dictionnaire	17
1.5.2	Les méthodes noyaux	18
<b>1.6</b>	<b>Combinaison de modèles</b>	18
<b>1.7</b>	<b>Conclusion</b>	19

### Chapitre 2. Méthodes statistiques de discrimination

GILBERT SAPORTA

<b>2.1</b>	<b>Introduction</b>	20
<b>2.2</b>	<b>Méthodes géométriques de discrimination sur données numériques</b>	20
2.2.1	Données et notation	20
2.2.2	Axes discriminants	22
2.2.3	Règles géométriques d'affectation	23
2.2.4	Insuffisance des règles géométriques	24
<b>2.3</b>	<b>Une méthode de discrimination sur variables qualitatives.</b>	
	<b>La méthode DISQUAL</b>	25
<b>2.4</b>	<b>Méthodes probabilistes</b>	26
2.4.1	La règle bayésienne	26
2.4.2	Méthodes non paramétriques	26
2.4.3	Le modèle normal multidimensionnel	27
<b>2.5</b>	<b>Mesures d'efficacité des règles de classement</b>	28

2.5.1	Taux d'erreur théorique pour deux groupes avec et distribution normale	28
2.5.2	La méthode de resubstitution	29
2.5.3	Les méthodes de validation croisée	29
2.5.4	Conclusion	30

### Chapitre 3. Méthodes neuronales de discrimination

PATRICK GALLINARI

3.1	<b>Introduction</b>	31
3.2	<b>Quelques modèles de réseaux</b>	32
3.2.1	Perceptrons multicouches et réseaux à fonctions radiales	32
3.2.2	Learning Vector Quantization	34
3.2.3	Arbres de neurones	35
3.2.4	Analogie structurelle entre modèles neuronaux et statistiques	36
3.3	<b>Critères d'apprentissage</b>	38
3.4	<b>Algorithmes</b>	40
3.5	<b>Interprétation probabiliste des sorties calculées par les réseaux</b>	43
3.5.1	Résultats généraux	43
3.5.2	Autres critères d'apprentissage	46
3.6	<b>Relations avec l'analyse discriminante</b>	47
3.7	<b>Estimation de densité</b>	49
3.8	<b>Conclusion</b>	51

### Chapitre 4. Détermination de l'architecture des réseaux multicouches. Application à un problème de discrimination : l'analyse automatique du sommeil

REGIS LENGELLE, NICOLAS SCHALTENBRAND

4.1	<b>Introduction : présentation du problème</b>	53
4.2	<b>Sélection de modèle et détermination de l'architecture de PMC</b>	58
4.2.1	Sélection de modèle	58
4.2.2	Recherche de l'architecture optimale	60
4.2.3	Algorithme de Cascade Corrélation	62
4.2.4	Optimisation de la Représentation Interne	63
4.3	<b>Application à l'analyse automatique du sommeil chez l'homme</b>	68
4.3.1	Présentation du problème	68
4.3.2	Description du système d'analyse automatique	70
4.3.3	Données d'apprentissage	70
4.3.4	Données test	70
4.3.5	Résultats obtenus	72
4.4	<b>Conclusion</b>	74

## Chapitre 5. Application d'un système de réseaux multicouches en télédétection : calcul des caractéristiques du vent en fonction des mesures radar

SYLVIE THIRIA, CARLOS MEJIA, FOUAD BADRAN, MICHEL CREPON

<b>5.1</b>	<b>Introduction</b>	75
<b>5.2</b>	<b>Le problème géophysique</b>	75
<b>5.3</b>	<b>La méthode</b>	78
5.3.1	Décodage du module : premier niveau	80
5.3.2	Décodage de la direction : deuxième niveau	81
<b>5.4</b>	<b>Les données</b>	83
<b>5.5</b>	<b>Performances</b>	84
5.5.1	Décodage de la vitesse : premier niveau	84
5.5.2	Décodage de la direction : deuxième niveau	85
<b>5.6</b>	<b>Conclusion</b>	88
<b>5.7</b>	<b>Annexe : Réseaux multicouches et classification</b>	88

## Partie 2 : Préviation

### Chapitre 6. Séries chronologiques. Les outils du statisticien, les besoins du prévisionniste

BERNARD BURTSCHY

<b>6.1</b>	<b>Introduction</b>	90
<b>6.2</b>	<b>Concepts de base</b>	90
6.2.1	Les séries chronologiques	91
6.2.2	Les composantes d'une série chronologique	91
6.2.3	Horizon prévisionnel	94
6.2.4	Les variables et leur stabilité	96
6.2.5	Le but de l'analyse d'une série chronologique	97
6.2.6	Niveau de complexité	99
<b>6.3</b>	<b>Décomposition d'une série chronologique</b>	99
6.3.1	Détermination d'une tendance	100
6.3.2	Déterminer un cycle	102
6.3.3	Déterminer la saisonnalité	102
<b>6.4</b>	<b>Le lissage exponentiel</b>	106
6.4.1	Le lissage exponentiel constant	106
6.4.2	Les diverses méthodes du lissage exponentiel	107
6.4.3	Lissage exponentiel à tendance linéaire	108
6.4.4	Lissage exponentiel à ajustement quadratique	109
6.4.5	Méthode de Holt-Winters	110
6.4.6	Automatisation du lissage exponentiel	110
<b>6.5</b>	<b>La méthode de Box-Jenkins univariée</b>	111
6.5.1	Stationnarité	111
6.5.2	Les modèles autorégressifs	111

6.5.3	Les modèles moyennes mobiles	112
6.5.4	Les modèles mixtes	113
6.5.5	Le théorème de Wold	114
6.5.6	Les étapes de la recherche de modèles ARMA	115
6.5.7	Rendre stationnaire une série	115
6.6	<b>Quelques éléments sur les méthodes explicatives</b>	116
6.7	<b>Conclusion</b>	118

## Chapitre 7. Une application des réseaux de neurones pour la prévision à un pas de temps

STEPHANE CANU, XIA DING, YVES GRANVALET

7.1	<b>Introduction</b>	120
7.2	<b>La prévision de consommation : un problème</b>	122
7.3	<b>Méthode utilisée</b>	127
7.3.1	Sélection des entrées	127
7.3.2	Prétraitements	127
7.3.3	Identification des paramètres	128
7.4	<b>Résultats</b>	128
7.5	<b>Les applications</b>	130
7.6	<b>Conclusion</b>	131

## Partie 3 : Représentation structurée

### Chapitre 8. Méthodes factorielles

LUDOVIC LEBART

8.1	<b>Introduction</b>	132
8.2	<b>Principes de base des méthodes factorielles</b>	132
8.2.1	Représentation élémentaire d'un tableau de données	132
8.2.2	Principales méthodes	134
8.2.3	Analyse générale : décomposition aux valeurs singulières	134
8.3	<b>Les principales méthodes</b>	140
8.3.1	L'analyse en composantes principales	140
8.3.2	L'analyse des correspondances	143
8.3.3	L'analyse des correspondances multiples	144
8.4	<b>Liens avec les réseaux de neurones</b>	146
8.4.1	Un Perceptron multicouche particulier	147
8.4.2	Un Perceptron multi-couche non-supervisé	149
8.4.3	Un réseau linéaire adaptatif	150

## Chapitre 9. Analyse de données multidimensionnelles par réseaux de neurones auto-organisés

JEANNY HERAULT, ANNE GUERIN-DUGUE

<b>9.1</b>	<b>Données et processus</b>	152
<b>9.2</b>	<b>Réduction de données et recherche de structure</b>	153
9.2.1	Petits échantillons et grandes dimensions	153
9.2.2	Phénomène de l'espace vide	154
9.2.3	Réduction de dimension	156
9.2.4	Réduction du nombre d'échantillons	157
9.2.5	Recherche de structure	158
<b>9.3</b>	<b>Représentation des données par cartes de Kohonen</b>	161
9.3.1	Principe	161
9.3.2	Exemple d'application en grandes dimensions	163
9.3.3	Intérêt et limitations de l'algorithme de Kohonen	163
<b>9.4</b>	<b>L'algorithme ACC</b>	164
9.4.1	Principe	164
9.4.2	Recherche de la projection	165
9.4.3	Exemple de résultat	167
9.4.4	Qualité de la représentation topologique	167
9.4.5	Interpolation et extrapolation	168
9.4.6	Projections inverses	169
<b>9.5</b>	<b>Conclusion</b>	170

## Chapitre 10. Classification non supervisée

YVES LECHEVALLIER

<b>10.1</b>	<b>Introduction</b>	171
10.1.1	Notations	172
<b>10.2</b>	<b>Modèles de mélange</b>	175
<b>10.3</b>	<b>Méthodes de partitionnement</b>	177
10.3.1	Algorithme itératif général	178
10.3.2	Algorithme de voisinage	178
10.3.3	Algorithmes possédant deux étapes d'optimisation	180
10.3.4	Fiabilité d'un algorithme de voisinage	184
10.3.5	Choix du nombre de classes	184
<b>10.4</b>	<b>Méthodes hiérarchiques</b>	184
10.4.1	Hiérarchie et distance ultramétrique	184
10.4.2	Algorithme général de la classification ascendante hiérarchique	186
10.4.3	Les mesures d'agrégation entre les classes	187
10.4.4	Construction de la hiérarchie indicée	187
10.4.5	La formule de récurrence de Lance et de Williams	188
<b>10.5</b>	<b>Méthodes basées sur la théorie des graphes</b>	188
10.5.1	Introduction	188
10.5.2	Arbre de longueur minimum	189
<b>10.6</b>	<b>Conclusion</b>	189

**Chapitre 11. Cartes topologiques et nuées dynamiques****FATIHA ANOUAR, FOUAD BADRAN, SYLVIE THIRIA**

<b>11.1</b>	<b>Introduction</b>	190
<b>11.2</b>	<b>L'algorithme des cartes topologiques</b>	190
<b>11.3</b>	<b>Cartes topologiques et nuées dynamiques</b>	193
11.3.1	Définitions et notations	193
11.3.2	Algorithme de Kohonen révisé	193
11.3.3	Version nuées dynamiques des cartes topologiques	194
11.3.4	Analogie avec l'algorithme de Kohonen	196
<b>11.4</b>	<b>Carte topologique et approximation de la fonction densité</b>	197
11.4.1	Formalisme probabiliste des cartes topologiques	197
11.4.2	Estimation des paramètres	199
11.4.3	Analogie avec l'algorithme de Kohonen	201
<b>11.5</b>	<b>Expérimentations</b>	203
<b>11.6</b>	<b>Conclusion</b>	206

**Chapitre 12. Régression par carte topologique****FOUAD BADRAN, PHILIPPE DAIGREMONT, SYLVIE THIRIA**

<b>12.1</b>	<b>Introduction</b>	207
<b>12.2</b>	<b>Régression par fonctions noyaux</b>	208
<b>12.3</b>	<b>L'approche par discrétisation</b>	209
<b>12.4</b>	<b>Régression par carte topologique</b>	211
12.4.1	L'algorithme	211
12.4.2	Mise en œuvre de SOM-3	214
<b>12.5</b>	<b>Analyse de la régression par carte topologique</b>	214
<b>12.6</b>	<b>expériences :</b>	216
12.6.1	Données simulées	216
12.6.2	Données réelles	218
<b>12.7</b>	<b>Conclusion</b>	221

**Chapitre 13. Cartes de kohonen : applications****FIONN MURTAGH**

<b>13.1</b>	<b>Introduction</b>	223
<b>13.2</b>	<b>La méthode SOFM de Kohonen</b>	224
<b>13.3</b>	<b>Classification sous contrainte de contiguïté (CCC)</b>	225
<b>13.4</b>	<b>Application de la méthode CCC</b>	226
<b>13.5</b>	<b>L'interprétation des résultats en utilisant les traceurs</b>	228
<b>13.6</b>	<b>Conclusion</b>	229

## Partie 4 : Validation, validité

### Chapitre 14. Heuristiques pour la généralisation

PATRICK GALLINARI

14.1	Introduction	230
14.2	Le compromis biais/variance	231
14.3	Choix de l'architecture	234
14.4	Stopper l'apprentissage avant convergence	235
14.5	Régularisation	236
14.6	L'élagage	238
14.7	Bruiter les entrées	239
14.8	L'interprétation bayésienne	241
14.9	Discussion	243

### Chapitre 15. La dimension de Vapnik-Chervonenkis : application aux réseaux de neurones

OLIVIER GASCUEL

15.1	Introduction	244
15.2	Trois formes de convergence	245
15.2.1	Notations, définitions	245
15.2.2	Convergence faible	247
15.2.3	Convergence forte	247
15.2.4	Consistance	248
15.3	Convergence uniforme en probabilité	248
15.3.1	Apprentissage et convergence uniforme en probabilité	248
15.3.2	Conditions nécessaires et suffisantes pour la convergence uniforme	251
15.4	Dimension de Vapnik-Chervonenkis (VC)	252
15.4.1	Définitions	252
15.4.2	Exemples	253
15.4.3	Propriétés	254
15.5	Application au perceptron multicouches (PMC)	255
15.5.1	Rappels, notations	255
15.5.2	Apprentissage	257
15.5.3	Le cas des fonctions de transfert à seuil	259
15.5.4	Résultats dans le cas général	259
15.6	Conclusion	260